Class 5: Conditional Probability & Independence

Tuesday, September 17, 2019 11:36 AM

Fun article: https://www.newyorker.com/magazine/2019/09/09/what-statistics-can-and-cant-tell-us-about-ourselves

HW 2 due Thurs, 1pm, Gradescope

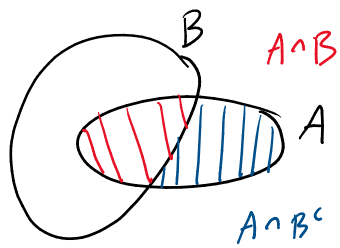
Recap:

- For events A and B, with $P(B) > 0$, define

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

given

- Multiplication rule: $P(A \cap B) = P(A \mid B) P(B)$
- Bayes Rule: $P(B \mid A) = \dfrac{P(A \mid B) P(B)}{P(A)}$.
- Law of total probability:



$A \cap B$ and $A \cap B^c$ are disjoint,

and $(A \cap B) \cup (A \cap B^c) = A$,

so $P(A) = P(A \cap B) + P(A \cap B^c)$

$= P(A \mid B) P(B) + P(A \mid B^c) P(B^c)$ (multiplication rule)

Medical Diagnostic Tests:

Get a medical test, comes back positive.
What is the prob. I have the disease?

False positive rate: $q = P(B \mid A^c)$

False negative rate: $r = P(B^c \mid A)$

Two events: A = "I have disease"

B = "test positive for disease"

Prob I have disease given I test positive

$= P(A \mid B)$

Prob I have disease given I test positive

$$= P(A \mid B).$$

Prevalence: probability of having disease or fraction of population with disease $= P(A) = p$.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \qquad \text{(Bayes rule)}$$

We know $P(B|A) = 1 - P(B^c|A) = 1 - r$

and $P(A) = p$.

Use law of total probability to get

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c)$$

$$= \left(1 - P(B^c|A)\right)P(A) + P(B|A^c)\left(1 - P(A)\right)$$

$$= (1-r)p + q(1-p).$$

Thus $$P(A|B) = \frac{p(1-r)}{p(1-r) + q(1-p)}.$$

p = prevalence

q = false pos.

r = false neg.

Ex1 $p = q = r = 1/1000000$.

$$\hookrightarrow P(A|B) = \frac{p(1-p)}{p(1-p) + p(1-p)} = \frac{p(1-p)}{2p(1-p)} = \frac{1}{2}.$$

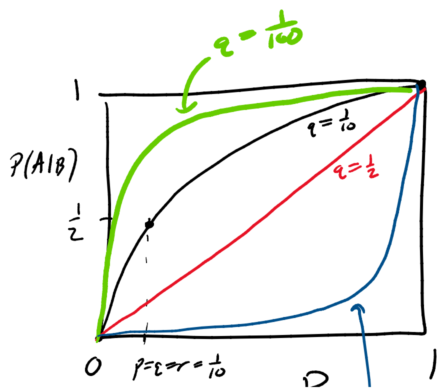
Ex2: Assume $r = q$. Then

$$P(A|B) = \frac{p(1-q)}{p(1-q) + q(1-p)} = \frac{p(1-q)}{p + q - 2pq}.$$

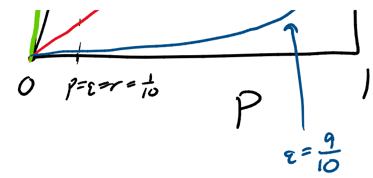
Ex 2a: If $q = \frac{1}{10} = r$,

$$P(A|B) = \frac{\frac{9}{10}p}{p + \frac{1}{10} - \frac{1}{5}p}.$$

As p decreases, test becomes less accurate!



As p decreases, test becomes less accurate!



Independence:

Recall $\quad P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$.

What does it mean for A and B to be independent?

Defn 1: Let A and B are events with $P(B) > 0$.
Then A and B are independent if

$$P(A \mid B) = P(A).$$

↳ Knowing B does not make A any more or less likely to occur.

↳ Only valid if $P(B) > 0$,

Defn 2 ("real"): Events A and B are independent if $\quad P(A \cap B) = P(A)P(B)$

↳ Note $P(A \cap B) = P(A \mid B)P(B)$ if $P(B) > 0$,
so equivalent to Defn 1 if $P(B) > 0$.

↳ Independence is symmetric:
"A indep. of B" is the same as "B indep. of A"

Ex 1: 1 in 10 shishito peppers are spicy.
I draw two from a bag. $\quad \Omega = \{Y, N\}^2$

yes spicy $\quad$ no not spicy

Let A = "first one is spicy" $= \{YY, YN\}$
B = "second one is spicy" $= \{NY, YY\}$
C = "not both spicy" $= \{NN, NY, YN\}$.

Probabilities:
$$P(YY) = \frac{1}{10} \cdot \frac{1}{10} = \frac{1}{100}$$
$$P(YN) = P(NY) = \frac{1}{10} \cdot \frac{9}{10} = \frac{9}{100}$$

$$P(YN) = P(NY) = \frac{1}{10} \cdot \frac{9}{10} = \frac{9}{100}$$
$$P(NN) = \frac{9}{10} \cdot \frac{9}{10} = \frac{81}{100}.$$

Q1: Are A and B independent?

$$P(A \cap B) = P(\{YY\}) = \frac{1}{100}$$
$$P(A)P(B) = \frac{1}{10} \cdot \frac{1}{10} = \frac{1}{100}$$
$$\to P(A) = P(\{YY, YN\}) = P(YY) + P(YN) = \frac{1}{100} + \frac{9}{100} = \frac{1}{10}.$$

Q2: Are A and C independent?

$$P(C) = P(NN) + P(NY) + P(YN) = \frac{81+9+9}{100} = \frac{99}{100}$$

$$P(A \cap C) = P(YN) = \frac{9}{100}$$
$$P(A)P(C) = \frac{1}{10} \cdot \frac{99}{100} = \frac{99}{1000}$$

Ex 2: Roll 2 dice. $\Omega = \{1,2,3,4,5,6\}^2$.

All outcomes have equal prob. $\frac{1}{36}$.

Let $A =$ "1st roll is 3" $= \{31, 32, 33, 34, 35, 36\} \to P(A) = \frac{6}{36} = \frac{1}{6}$

$B =$ "sum is 7" $= \{16, 25, 34, 43, 52, 61\} \to P(B) = \frac{6}{36} = \frac{1}{6}$

Are A and B independent? YES.

$$A \cap B = \{34\} \to P(A \cap B) = \frac{1}{36}$$
$$P(A)P(B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

Defn: Let $E_1, E_2, \dots, E_n$ be events.

- They are pairwise independent if $P(E_i \cap E_j) = P(E_i)P(E_j)$ for all $i \neq j$.

- They are (jointly) independent if

$$P(E_{i_1} \cap E_{i_2} \cap \dots \cap E_{i_k}) = P(E_{i_1})P(E_{i_2}) \dots P(E_{i_k})$$

for $1 \le \underbrace{i_1 < i_2 < \cdots < i_k}_{\text{which ones to look at}} \le n$ for $\underset{\substack{\uparrow \\ \text{how many of } E_i\text{'s} \\ \text{to look at}}}{1 \le k \le n}$.

<u>Ex</u>: <u>$n=3$</u>. Events A, B, and C are <u>pairwise</u> <u>indep.</u> if

$$P(A \cap B) = P(A)P(B)$$
$$P(A \cap C) = P(A)P(C)$$
$$P(B \cap C) = P(B)P(C).$$

They are <u>jointly</u> <u>independent</u> if also $P(A \cap B \cap C) = P(A)P(B)P(C)$.

<u>Fact</u>: Joint indep. implies pairwise indep.
Pairwise indep. does <u>not</u> imply joint indep. $\longrightarrow$ HW 3

<u>Intuition</u>: A, B, C can be pairwise indep., but still knowledge of C provides info about "dependence" between A and B.

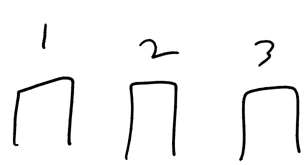
$$\boxed{\begin{matrix} A \text{ indep } B \\ \text{and} \\ B \text{ indep } C \end{matrix} \not\Rightarrow A \text{ indep } C}$$

<u>Ex</u>: $\Omega = \{H, T\}^n$ n coin flips, each outcome equally likely $\rightarrow \frac{1}{2^n}$

Let E_i = "i^{th} flip is heads".

<u>Check</u>: $E_1, E_2, \ldots, E_n$ are <u>jointly</u> independent.

<u>Monty Hall Problem</u>:



<u>Game Show</u>: Pick 1 door of 3,

- One door hides a car (prize),
- Other two hide goats.
- After you pick a door, Monty opens one of the other 2 doors to reveal a goat.

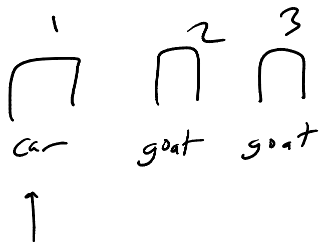
opens one of the other 2 doors to reveal a goat.

- If both other doors are goats, he opens one at random.
- Then you get to switch your choice if you want! → Should you switch?
  ↳ YES

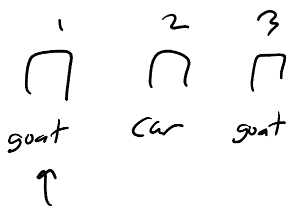
- If you switch, the prob of winning is $2/3$, otherwise $1/3$.

- SWITCH: Pick door #1.

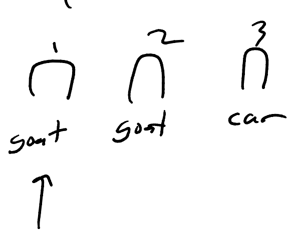
  3 cases:

(1) 

→ I lose

(2) 

→ I win

(3) 

→ I win